心理学报 2023, Vol. 55, No. 1, 1−8
Acta Psychologica Sinica

https://doi.org/10.3724/SP.J.1041.2023.00001

·工程心理学专栏(Section of Engineering Psychology)·

听觉刺激对虚拟环境中空间压缩的影响*

胡晓斐 汪嘉维 刘涵宇 宋晓蕾

(陕西师范大学心理学院; 陕西省行为与认知神经科学重点实验室, 西安 710062)

摘 要 人们往往会低估虚拟环境中空间的距离(空间压缩)。早期研究发现视听刺激的不一致会改善空间压缩, 但听觉刺激在其中的作用及其影响效果需进一步研究。本研究首先探究了听觉刺激的有无对空间压缩的影响, 接着操控视听刺激之间的物物距离进一步探究其对空间压缩的影响和改善。研究发现听觉刺激以两种方式影响空间压缩: 听觉刺激的存在可以改善空间压缩; 空间压缩与视听刺激的物物距离负相关。研究建议为了改善空间压缩, 视听刺激应当同时呈现且物物距离至少为 1 m。

关键词 虚拟现实, 空间压缩, 视听刺激, 物我距离, 物物距离

分类号 B842

1 背景介绍

虚拟现实(Virtual Reality, VR)作为一项新兴技术, 在科研、制造、军事、技能培训等诸多领域均有着越来越广泛的应用。例如, 消防员可以在虚拟火灾场景中进行多次训练, 从而起到降低成本和避免实操伴随的潜在危险的作用(Saghafian et al., 2020); 在心理学领域, 可以通过 VR 制作各种刺激和任务场景, 提高心理学实验的生态有效性和可重复性(Pan & Hamilton, 2018)。然而, VR 的普及仍然面临着严峻的挑战, 如通信延迟、渲染质量不高及空间压缩等。虽然随着 5G 网络和 3D 引擎技术的发展, 硬件和软件相关的挑战逐渐得到解决, 但由于用户自身所产生的挑战(如空间压缩)仍然没有得到足够的重视。

距离知觉就是由于用户自身所产生的挑战之一。它代表着受试者感知两个位置的空间距离的能力, 可以分为两种类型: 当受试者自己被视为一个位置时, 自己与另外位置的距离被称为物我距离; 当两个位置均在受试者的外部时, 两个位置的距离被称为物物距离。早在 20 世纪, 心理学家们就发现了 VR 中的空间压缩, 即人们经常低估虚拟环境中的空间距离(Henry & Furness, 1993; Interrante et al., 2008; Naceri et al., 2009; Renner et al., 2013)。这种现象对虚拟环境中的物我距离感知和物物距离感知都是成立的, 因此极大的阻碍了 VR 的应用, 特别是需要精细操作的场景(Li et al., 2011)。虚拟和现实环境之间明显的感知差异破坏了用户的沉浸式体验, 降低了他们的接受度。因此, 改善空间压缩以提高保真度并最终促进 VR 的广泛应用是至关重要的。

研究者们已经报道了几种改善空间压缩的方法。例如, Kelly 等人(2014)发现, 在盲走实验中, 与没有视觉反馈的条件相比, 当视觉反馈存在时受试者对于距离的判断能力得到了极大的改善。然而, 这种改善的效果是边际递减的。视觉目标的存在只对前五次试验的距离判断有明显的改善, 而在接下来的试验中, 改善程度就会下降。Interrante 等人(2006)发现, 当受试者处于一个与现实环境相似的虚拟环境中时, 空间压缩在很大程度上被削弱了。

收稿日期: 2022-04-15

* 国家自然科学基金面上项目(32071065); 国家自然科学基金重大项目(T2192931); 载人航天工程航天医学实验领域项目(HYZHXM03001); 中央高校基本科研业务费重点项目(GK202002010)。

通信作者: 宋晓蕾, E-mail: songxiaolei@snnu.edu.cn

然而，这样的改善会失去 VR 的最大优势，即创造一个在现实环境中很少经历的虚拟环境。此外，空间压缩是一种多模态现象，可以发生在听觉和视觉刺激上(Rébillat et al., 2012)。Huang 等人(2021)使用音频混响来改善受试者的空间压缩。Finnegan 等人(2016)发现，当虚拟环境中的听觉和视觉信息在空间上不一致时，受试者的距离判断可以更准确。这里的不一致意味着听觉和视觉刺激之间的物物距离不为零。虽然他们没有进一步研究不一致程度对于空间压缩的改善效果，但听觉和视觉刺激的结合是有潜力起到改善作用的。

当人们从两种或更多的模态中接受信息时，他们会整合所有模态的信息，然后得到一个最终的感知判断(Alais & Burr, 2004; Ernst & Banks, 2002; Finnegan et al., 2016)。以听觉和视觉模态下的距离感知为例，受试者的最终判断将是听觉和视觉刺激引起的距离感知的加权总和。如果在虚拟环境中，听觉刺激比视觉刺激相对于受试者呈现得更远(即听觉刺激的物我距离更大)，那么对于视觉刺激的空间压缩会由于听觉刺激的存在而得到改善。然而，这样的多模态整合理论仅适用于对一个物体的感知。研究表明，听觉和视觉刺激需要保持在一定范围内才能被感知为一个物体，即所谓的腹语效应(Bruns, 2019)。因此，在使用不一致的视听刺激来改善空间压缩时，应当考虑听觉和视觉刺激之间的物物距离。

此外，受试者和视觉刺激之间的物我距离也很重要。Finnegan 等人(2016)强调，当视觉刺激与受试者的物我距离很小时(在他们的研究中为 3 m)，不一致的视听刺激并不能改善空间压缩。他们认为改善的失败是由于受试者会在物我距离较小时高估距离而不是低估距离。虽然研究者们都同意人们会高估离自己近，低估离自己远的刺激的距离，但是对于发生高估和低估现象的转折点的距离并没有统一的答案(在不同的研究中分别为 1 m、1.9 m 或 3.23 m) (Anderson & Zahorik, 2014; Armbrüster et al., 2008; Finnegan et al., 2016; Zahorik et al., 2005)。尽管如此，受试者和视觉刺激之间的物我距离足够大时，就可以很稳定地引起空间压缩。

以往对于虚拟现实中空间压缩的研究一般采用四种方法：口头报告法，感知引导行为，想象行为和感知匹配法(Klein et al., 2009; Maruhn et al., 2019; Rébillat et al., 2012)。在口头报告法中，受试者需要以米或英尺等距离计量单位口头回答自己感知到的距离，但是该方法易受到受试者早期生活经验中对于距离计量单位认知的影响。在感知引导行为法中，受试者需要通过盲走或三角盲走等动作，将自己所走过的实际距离作为自己对最初看到的目标距离的回答，但是该方法对于实验场地的要求较高。在想象行为法中，受试者需要在大脑中想象自己的行走，通过将想象行走的时间与受试者的行走速度相乘，得到其对于当前目标的距离判断，但是该方法会受到受试者对于时间长度认知的影响。在感知匹配法中，受试者可利用心理物理学中的调整法，不断调节刺激的位置使其与之前出现的目标刺激距离一致，抑或是心理物理学中的固定刺激法，多次进行二选一的强制回答任务，得到最终的距离判断(Ellis & Menges, 1998; Wu et al., 2004)。感知匹配法利用认知心理学领域中的心理物理学方法，避免了以上三种方式的缺点，通过测量受试者的主观相等点，得到最终的距离判断结果。尽管如此，以往研究所采用的调整法和固定刺激法仍有缺点。调整法易受到受试者适应效果的影响，对于绝对阈值的测量并不准确；固定刺激法则极其花费时间，同时对于刺激相关参数的选择要求极高。鉴于此，本研究采用心理物理学适应法中的阶梯法，通过受试者对上一个试次中距离判断任务的二选一强制性回答，实时调整下一试次中视觉刺激的相关参数(本研究中为调整视觉刺激的物我距离)，从而在减少实验时间的同时，有效地测量每位受试者的主观距离判断。

在本研究中，我们进行了两个实验来研究听觉刺激对 VR 中空间压缩的影响和改善。尽管与一致的视听刺激相比，不一致的视听刺激可以改善空间压缩，但是我们不清楚一致的视听刺激是否会比纯视觉刺激表现得更好(Finnegan et al., 2016)。因此，在实验 1 中，我们操控了听觉刺激是否存在。此外，我们还改变了视觉刺激的物我距离(3 m、4 m、5 m)。其中，最小的距离(3 m)超过了先前研究所报道的大部分转折点的距离，确保了空间压缩在本研究实验过程中的出现(Anderson & Zahorik, 2014; Armbrüster et al., 2008; Finnegan et al., 2016; Zahorik et al., 2005)。此外，以往研究大都测量了物我距离为 5 m 时的空间压缩现象(Armbrüster et al., 2008; Buck et al., 2021; Finnegan et al., 2016; Rébillat et al., 2012)。为了与以往研究的结果相对比，我们也设置了物我距离为 4 m 和 5 m 时的实验条件。然后，在实验 2 中，我们操控了视听条件的

不一致程度, 即听觉和视觉刺激的物物距离(0.5 m、1 m、1.5 m、2 m)。虽然 Gorzel 等人(2012)发现腹语效应的可接受的物物距离会随着视觉刺激的物我距离的增加而增加, 但 3 m 和 5 m 之间的微小差异应该不会对可接受的物物距离产生任何影响。

2 实验 1：听觉刺激存在与否对空间压缩的影响

2.1 目的

早前研究发现, 与一致的视听刺激相比, 不一致的视听刺激可以改善空间压缩, 但一致的视听刺激是否优于纯视觉刺激还不得而知(Finnegan et al., 2016)。因此, 我们在这个实验中操控了听觉刺激的存在与否来考察其对空间压缩的影响。

2.2 被试

30 名大学生(13 名男性和 17 名女性, 17~26 岁, 平均年龄 ＝ 20.3 岁)被招募参加实验 1 和实验 2。实验 1 和实验 2 在同一天完成。为了避免练习效应, 对于每个受试者来说, 实验 1 和实验 2 的全部条件均随机呈现。每位受试者都有正常或矫正后的正常视力, 并且不具有脑部或精神疾病的历史, 我们也征得了每位受试者的书面同意。本研究的实验方案得到了陕西师范大学伦理委员会的批准。

2.3 实验设备与刺激材料

我们使用HTC Vive Pro来渲染虚拟环境, 内置的耳机被用来呈现听觉刺激, 一个键盘放在手边, 以接收受试者的反应。

VR 场景是用虚幻引擎(4.25 版)制作的。我们使用了引擎的内部材料：地面是白色, 环境完全开放; 自然阳光被用来作为环境光, 从受试者的背后照射; 视觉刺激是由非反射材料组成。视觉刺激是两个 1 m × 1 m 的方块, 其中心距地面 1.5 m。受试者的眼睛与视觉刺激保持水平, 以避免视觉水平线对距离感知的影响(Rand et al., 2011)。参考刺激是一个绿色方块, 有三种物我距离：3 m、4 m 和 5 m; 调整刺激是一个蓝色方块, 其初始的物我距离为比参考刺激的物我距离小 0.5 m。Armbrüster 等人(2008)在他们的实验里发现, 当物我距离为 4 m 时, 23 名受试者的平均主观距离为 3.2 ± 0.3 m 左右。为了使得实验中每一个试次中的视觉刺激的物我距离都在最终结果附近变化, 对于 4 m 物我距离的实验来说, 我们设置调整刺激的初始物我距离为 3.5 m (即比参考刺激小 0.5 m)。由于我们有三种物我距离的条件(3 m、4 m、5 m), 我们将比参考刺激小 0.5 m 作为每一个物我距离的初始设置。实验中所使用的噪音为粉红噪音(粉红噪音的频率能量分布主要在中低频段, 更容易被人耳所接受, 因此常用于声学测试), 噪音声源选取了音量大小全程无变化的声音, 强度为 60 dB, 时长为 5 秒。

2.4 实验程序

在实验 1 中, 听觉刺激有两种条件, 分为存在或不存在。受试者首先需要进行一个练习实验, 以了解实验说明并熟悉 HTC Vive Pro 的操作。练习实验包括 10 次试验, 受试者应通过按“A” (代表蓝色方块)或“D” (代表绿色方块)来回答哪个方块(绿色或蓝色)离他们更近。练习结束后, 受试者被要求进行距离判断任务。由于视觉反馈可以改善空间压缩, 从而污染听觉刺激的效果, 因此参考视觉刺激在一个实验中仅呈现一次(Kelly et al., 2014)。

每种条件的实验的第一次试验开始时有 5 秒的参考期间, 在此期间呈现参考视觉刺激(绿色方块); 如果当前实验包括听觉刺激, 则会有 5 秒的粉红噪声与参考视觉刺激同时呈现。接着, 参考视觉刺激和听觉刺激(如果存在的话)同时消失。0.5 秒后, 调整视觉刺激(蓝色方块)会被呈现。受试者需要判断在调整期间, 蓝色方块是否比最初呈现的绿色方块离自己更近。当受试者认为蓝色方块更近时, 应该按下“A”, 否则按下“D”。受试者一旦作出反应, 蓝色方块就会消失 0.5 秒, 下一次试验开始, 并呈现出一个新的蓝色方块。如果受试者在前一次试验中认为蓝色方块更近, 则当前试验新呈现蓝色方块的物我距离将向近手侧移动 5 cm, 反之向远手侧移动 5 cm。连续进行 30 次试验后, 当前实验结束。最后一次试验中蓝色方块的物我距离, 将作为受试者对初始绿色方块的最终距离判断。受试者应当在没有速度要求的前提下, 尽可能准确地回答距离判断任务。

对于每个实验, 参考刺激的物我距离(3 m、4 m 和 5 m)是固定的。此外, 在实验 1 中, 听觉刺激的存在或不存在在一个实验中也是固定的, 因此实验 1 总共有 6 个实验。

2.5 结果

我们对空间压缩率进行了 3×2 的重复测量方差分析。为了使得空间压缩率为正值以美化图示, 我们在参考 Armbrüster 等人(2008)的研究的基础上, 将空间压缩率定义为实际的物我距离减去受试者感知的物我距离, 然后除以实际的物我距离。被试内因素包括参考刺激的物我距离(3 m、4 m 和 5 m)

和听觉刺激的存在(视听条件和纯视觉条件)。

空间压缩率的结果如图 1 所示，其中横坐标代表参考刺激的物我距离，纵坐标代表空间压缩率，不同的颜色分别代表视听条件和纯视觉条件。我们发现，视听条件下的空间压缩率要小于纯视觉条件下的压缩率，$F(1, 29) = 4.05$, $p = 0.054$, $\eta_p^2 = 0.12$。然而，参考刺激的物我距离的主效应和两个被试内因素之间的交互作用并不显著($ps > 0.05$)。其描述性统计由表 1 所示。

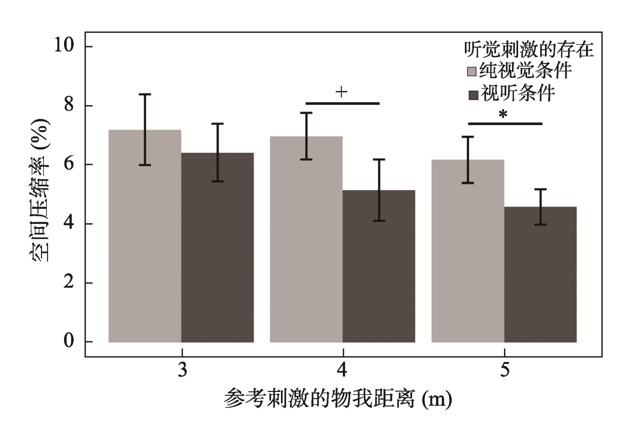


图 1 实验 1 空间压缩率的结果。横轴代表参考刺激的物我距离；纵轴代表空间压缩率。深色条和灰色条分别表示视听和纯视觉条件。可达到显著水平的事后统计结果呈现在相应的位置。黑色的误差线代表 30 名受试者的平均值的标准误差。

注：*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$。

表 1 实验 1 空间压缩率的描述性统计($M \pm SE$)

听觉刺激的存在	实际距离		
	3 m	4 m	5 m
纯视觉条件	7.22% ± 1.21%	7.00% ± 0.80%	6.20% ± 0.79%
视听条件	6.44% ± 0.98%	5.17% ± 1.04%	4.60% ± 0.60%
视听−纯视觉	$t = -0.51$,	$t = -1.59$,	$t = -1.85$, *

注：*表示 $p < 0.05$

然后，我们对在不同的参考刺激的物我距离下的视听条件和纯视觉条件的空间压缩率进行了配对的单侧 t 检验，自由度为 29，其结果如表 1 所示。我们发现，在物我距离为 4 m (差别 = 1.83%, $p = 0.062$)和 5 m (差别 = 1.60%, $p = 0.037$)时，视听条件的空间压缩率小于纯视觉条件，但物我距离为 3 m 时则不然($p = 0.307$)。

受试者在视听条件下能比纯视觉条件下做出更准确的距离判断，尤其是对于更远的视觉刺激(物我距离为 4 m 或 5 m 时)，这种效果就更加明显，可分别达到显著的 1.83%和 1.60%的效果提升。早期研究已经报道了听觉刺激可以用来扩展受试者对于立体图像的深度感知(Turner et al., 2011)。因此，听觉刺激本身可以改善视觉刺激的空间压缩。在下一个实验中，我们将控制听觉和视觉刺激之间的物物距离，以研究视听条件的不一致程度对于空间压缩的影响。

3 实验 2：听觉和视觉刺激之间的物物距离对空间压缩的影响

3.1 目的

我们发现听觉刺激本身可以改善空间压缩。此外，早期研究表明，不一致的视听刺激可以进一步改善空间压缩(Finnegan et al., 2016)。因此，在本实验中，我们控制了听觉和视觉刺激之间的物物距离，以研究视听条件的不一致程度对空间压缩的影响。

3.2 被试、实验设备与材料

被试、视觉和听觉材料均与实验 1 相同。

3.3 实验流程

实验 2 与实验 1 的实验流程相同。在实验 2 中，听觉刺激一直存在，且其与视觉刺激之间的物物距离是 4 种条件：0.5 m、1 m、1.5 m 和 2 m。实验 2 总共有 12 个实验。

3.4 结果

与实验 1 类似，我们对空间压缩率进行了 3×5 的重复测量方差分析。由于实验 1 和实验 2 招募了相同的受试者，实验 1 中视听条件的结果，即物物距离为 0 m，也被包含在当前的分析中。被试内因素包括参考刺激的物我距离(3 m、4 m 和 5 m)以及听觉和视觉刺激物之间的物物距离(0 m、0.5 m、1 m、1.5 m 和 2 m)。

本实验发现，听觉和视觉刺激之间的物物距离的主效应是显著的，$F(4, 116) = 8.29$, $p < 0.001$, $\eta_p^2 = 0.22$。然而，参考刺激的物我距离的主效应和两个被试内因素之间的交互作用并不显著($ps > 0.05$)。其描述性统计由表 2 所示。

表 2 实验 2 空间压缩率的描述性统计($M \pm SE$)

物物距离(m)	实际距离		
	3 m	4 m	5 m
0	6.44% ± 0.98%	5.17% ± 1.04%	4.60% ± 0.60%
0.5	4.44% ± 1.43%	4.33% ± 0.71%	3.67% ± 0.67%
1	3.44% ± 1.06%	3.25% ± 1.01%	2.47% ± 0.73%
1.5	2.00% ± 1.95%	2.25% ± 1.73%	1.13% ± 0.99%
2	0.78% ± 1.23%	0.58% ± 0.80%	0.67% ± 0.63%

表 3　实验 2 空间压缩率的事后检验结果

物物距离(m)	*Z* 值	*p* 值
0.5 ~ 0	−1.38	0.648
1 ~ 0	-2.56^{+}	0.078
1.5 ~ 0	-3.93^{***}	<0.001
2 ~ 0	-5.15^{***}	<0.001
1 ~ 0.5	−1.19	0.756
1.5 ~ 0.5	-2.58^{+}	0.077
2 ~ 0.5	-3.80^{**}	0.002
1.5 ~ 1	−1.37	0.645
2 ~ 1	-2.59^{+}	0.072
2 ~ 1.5	−1.22	0.740

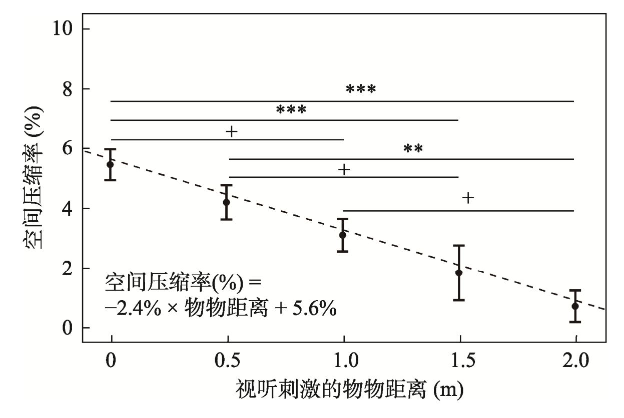


图 2　实验 2 空间压缩率的结果。横轴代表视听刺激的物物距离；纵轴代表空间压缩率。可达到显著水平的事后统计结果呈现在相应的位置。黑色的误差线代表 30 名受试者的平均值的标准误差。

注：*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$。

然后，在混合参考刺激的三个物我距离的结果之后，我们进行了 Tukey 多重比较检验，其结果如表 3 所示。空间压缩率的结果如图 2 所示，其中横坐标代表听觉和视觉刺激的物物距离，纵坐标代表空间压缩率。我们发现，空间压缩率随着物物距离的增加而降低。此外，当比较任意两个相距至少 1 m 的物物距离时，空间压缩率可以达到边缘显著或显著的差异($ps < 0.08$)。

最后，我们利用 R 语言的 lm 函数，拟合了一条以听觉和视觉刺激的物物距离(0 m、0.5 m、1 m、1.5 m 和 2 m)为自变量，空间压缩率为因变量的线性曲线(图 2 中的虚线)，结果如下：空间压缩率 = −2.4% × 物物距离 + 5.6%。其中，斜率可达到显著水平(p = 0.008)，表明空间压缩率与物物距离呈负相关，并且每 1 m 可以改善 2.4%的空间压缩率。

4　讨论

随着 5G 时代的到来，5G 网络很好地解决了通讯技术的时延、速率等瓶颈问题，但对 VR 技术的普及和应用提出了更高的要求，为了使该技术得到更好的推广与使用，虚拟环境中空间压缩这一瓶颈问题必须得到有效的改善或解决；此外，随着虚拟现实技术的成熟，其应用领域有望扩展到科研、制造、军队和消防员应急服务训练等多个领域，虚拟环境中正确的空间感知除了对 VR 技术某些功能有至关重要的作用外，还可以作为用户接受度的指标来衡量虚拟环境的保真度，因此，有效解决或改善该问题也是提高 VR 技术应用中以用户为中心的体验需求一个根本指标。基于此，本项目拟就虚拟环境中的空间压缩问题展开系统的研究，我们研究了听觉刺激对 VR 中空间压缩的影响和改善。

在实验 1 中，我们发现通过将听觉刺激与视觉刺激一起呈现，特别是当物我距离较大时，空间压缩可以得到改善。在实验 2 中，我们发现，如果听觉和视觉刺激之间的物物距离增加，空间压缩可以进一步改善。这是个简单的负相关关系。任何大于 1 m 的物物距离的增加都可以显著改善空间压缩。根据我们的结果，我们强烈建议在视觉刺激后面至少 1 m 处同时呈现听觉刺激以改善虚拟环境中的空间压缩。当物物距离的差距为 1 m 的时候，我们虽然得到的都是边缘显著的结果，但是我们仍将其认为是可以达到有效改善的。这是因为对于 1 m 和 1.5 m 之间更为细致的物物距离(1.1 m、1.2 m 等)，我们并没有进行实际测试。但是根据我们拟合的曲线(图 2)，即使是提升 0.1 m 的物物距离，空间压缩也会得到一定程度的改善，因此我们的建议是从物物距离为 1 m 开始，使得 VR 的设计师们有更大的选择。

我们在两个实验中都没有发现视觉刺激的物我距离的主效应，这与之前的研究并不一致(Buck et al., 2018; Buck et al., 2021)。我们将其归因于本研究中的物我距离之间的较小区别(最大区别为 2 m)。Buck 等人(2018)利用与本研究相同的设备(HTC Vive Pro)，测试了三种物我距离下(5 m、7.5 m 和 10 m)的空间压缩。他们虽然报告了物我距离的显著主效应，但 7.5 m 和 10 m 之间的差异并没有达到显著水平。因此，物我距离对空间压缩的影响并不稳定。

我们发现听觉刺激有两种影响空间压缩的方式。首先，与没有听觉刺激的情况相比，听觉刺激的存在可以提高距离判断的准确性。由于距离感知是一种多模态现象，听觉刺激所携带的额外信息可

以帮助受试者提高他们的最终判断的准确性(Rébillat et al., 2012)。其次，空间压缩率与听觉和视觉刺激之间的物物距离呈负相关。当受试者接受来自于一个物体的多种模态的信息时，他们倾向于计算每种模态的加权后的感知，然后整合它们以获得最终的判断(Berry et al., 2014; Ernst & Banks, 2002; Finnegan et al., 2016)。一旦听觉和视觉刺激的物我距离不一致，受试者的最终判断将是加权的听觉和视觉判断之和。以本文的实验条件为例，当听觉刺激呈现在视觉刺激后面时，听觉刺激携带的距离信息与视觉刺激携带的距离信息相比会更远，从而改善视觉刺激的空间压缩。

我们成功地为空间压缩率和物物距离拟合了一条线性曲线。固定的斜率系数(−0.024)表明，本研究中的听觉和视觉模态各自的权重保持不变。然而，权重是可以发生变化的。例如，通过改变视觉刺激的模糊程度，从而降低视觉模态的权重(Finnegan et al., 2016)。因此，尽管我们发现空间压缩可以通过每 1 m 物物距离的增加而实现 2.4%的改善，但改善的效果并不固定。如果通过某些操作增加了听觉模态的权重，如前面提到的模糊视觉刺激，改善的效果可能会增强。值得注意的是，物物距离不应该过大，否则腹语效应就会失灵，从而无法整合多种模态的信息(Bruns, 2019)。此外，不合适的物物距离也可能会导致距离的高估，而不是空间压缩的改善(Finnegan et al., 2016)。如图 2 所示，当物物距离为 2 m 时，空间压缩率约为 0.7%。因此，在本研究中，物物距离大于 2 m 时，可能会导致距离高估。

本研究仍存在局限性。首先，我们使用了一个修订的心理物理学范式。与典型的二选一方法不同，本研究中参考视觉刺激仅呈现一次，以消除视觉反馈对空间压缩的影响(Kelly et al., 2014)。然而，这样的操作可能会引起记忆问题。受试者必须记住参考刺激的感知距离，以便在一个实验内进行后续判断。我们需要在未来采用一种可以同时避免视觉反馈和记忆干扰的新范式。其次，本研究所采用的视觉刺激的物我距离(3 m、4 m 和 5 m)比较小，使得我们无法得到物我距离的主效应。我们需要在物我距离明显影响空间压缩的情况下进行实验，以进一步研究听觉刺激的效果。最后，我们没有进行纯听觉的实验。因此，在本研究中，我们无法计算出听觉和视觉模态各自的权重。我们需要在未来进行完整的实验(纯听觉、纯视觉和视听觉)。

5 结论

总而言之，我们发现听觉刺激可以通过两种方式影响空间压缩，与单纯的视觉条件相比，听觉和视觉刺激的同时存在可以改善空间压缩。此外，空间压缩率与听觉和视觉刺激之间的物物距离呈负相关。我们建议听觉和视觉刺激同时呈现，且之间的物物距离至少为 1 m。

参 考 文 献

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257–262.

Anderson, P. W., & Zahorik, P. (2014). Auditory/visual distance estimation: Accuracy and variability. *Frontiers in Psychology*, *5*, 1097. https://doi.org/10.3389/fpsyg.2014.01097

Armbrüster, C., Wolter, M., Kuhlen, T., Spijkers, W., & Fimm, B. (2008). Depth perception in virtual reality: Distance estimations in peri-and extrapersonal space. *Cyberpsychology & Behavior, 11*(1), 9–15.

Berry, J. S., Roberts, D. A., & Holliman, N. S. (2014, February). 3D sound and 3D image interactions: A review of audio-visual depth perception. In B. E. Rogowitz, T. N. Pappas, & H. D. Ridder (Eds.), *Proceedings of SPIE - The International Society for Optical Engineering: Vol. 9014: Human Vision and Electronic Imaging XIX* (pp.33–48). SPIE.

Bruns, P. (2019). The ventriloquist illusion as a tool to study multisensory processing: An update. *Frontiers in Integrative Neuroscience*, *13*, 51. https://doi.org/10.3389/fnint.2019.00051

Buck, L., Paris, R., & Bodenheimer, B. (2021). Distance compression in the htc vive pro: A quick revisitation of resolution. *Frontiers in Virtual Reality, 2,* Article 728667. https://doi.org/10.3389/frvir.2021.728667

Buck, L. E., Young, M. K., & Bodenheimer, B. (2018). A comparison of distance estimation in HMD-Based virtual environments with different HMD-Based conditions. *ACM Transactions on Applied Perception*, *15*(3), 1–15.

Ellis, S. R., & Menges, B. M. (1998). Localization of virtual objects in the near visual field. *Human Factors, 40*(3), 415–431.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

Finnegan, D. J., O'Neill, E., & Proulx, M. J. (2016, May). Compensating for distance compression in audiovisual virtual environments using incongruence. In K. Jofish, & D. Allison (Eds.), *CHI'16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 200–212). Association for Computing Machinery.

Gorzel, M., Corrigan, D., Squires, J., Boland, F., & Kearney, G. (2012, March). *Distance perception in virtual audio-visual environments*. Paper presented at UK 25th Conference: Spatial Audio in Today's 3D World. York, UK. http://www.aes.org/e-lib/browse.cfm?elib=18119

Henry, D., & Furness, T. (1993, September). *Spatial perception in virtual environments: Evaluating an architectural application*. Paper presented at IEEE Virtual Reality Annual International

Symposium. Seattle, Washington, USA: IEEE. https://doi.org/10.1109/VRAIS.1993.380801

Huang, Y. H., Venkatakrishnan, R., Venkatakrishnan, R., Babu, S. V., & Lin, W. C. (2021, September). *Using audio reverberation to compensate distance compression in virtual reality.* Paper presented at SAP '21: ACM Symposium on Applied Perception 2021, New York, USA: ACM, 1–10. https://doi.org/10.1145/3474451.3476236

Interrante, V., Ries, B., & Anderson, L. (2006, March). *Distance perception in immersive virtual environments, revisited.* Paper presented at IEEE Virtual Reality Conference (VR 2006). Alexandria, Virginia, USA: IEEE. https://doi.org/10.1109/VR.2006.52

Interrante, V., Ries, B., Lindquist, J., Kaeding, M., & Anderson, L. (2008). Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments. *Presence: Teleoperators and Virtual Environments, 17*(2), 176–198.

Kelly, J. W., Hammel, W. W., Siegel, Z. D., & Sjolund, L. A. (2014). Recalibration of perceived distance in virtual environments occurs rapidly and transfers asymmetrically across scale. *IEEE Transactions on Visualization and Computer Graphics*, *20*(4), 588–595.

Klein, E., Swan, J. E., Schmidt, G. S., Livingston, M. A., & Staadt, O. G. (2009, March). *Measurement protocols for Medium-Field distance perception in Large-Screen immersive displays*. Paper presented at 2009 IEEE Virtual Reality Conference, Lafayette, Louisiana, USA: IEEE. https://doi.org/10.1109/VR.2009.4811007

Li, Z., Phillips, J., & Durgin, F. H. (2011). The underestimation of egocentric distance: Evidence from frontal matching tasks. *Attention, Perception & Psychophysics*, *73*(7), 2205–2217.

Maruhn, P., Schneider, S., & Bengler, K. (2019). Measuring egocentric distance perception in virtual reality: Influence of methodologies, locomotion and translation gains. *PloS One*, *14*(10), Article e0224651. https://doi.org/10.1371/jounal.pone.0224651

Naceri, A., Chellali, R., Dionnet, F., & Toma, S. (2009, November). *Depth perception within virtual environments: A comparative study between wide screen stereoscopic displays and head mounted devices*. Paper presented at 2009 Computation World: Future Computing, Service Computation, Cognitive, Adaptive, Content, Patterns (COMPUTATIONWORLD). Athens, Greece: IEEE. https://doi.org/10.1109/ComputationWorld.2009.91

Pan, X., & Hamilton, A. F. de C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), 395–417.

Rand, K. M., Tarampi, M. R., Creem-Regehr, S. H., & Thompson, W. B. (2011). The importance of a visual horizon for distance judgments under severely degraded vision. *Perception*, *40*(2), 143–154.

Rébillat, M., Boutillon, X., Corteel, É., & Katz, B. F. (2012). Audio, visual, and audio-visual egocentric distance perception by moving subjects in virtual environments. *ACM Transactions on Applied Perception*, *9*(4), 1–17.

Renner, R. S., Velichkovsky, B. M., & Helmert, J. R. (2013). The perception of egocentric distances in virtual environments-a review. *ACM Computing Surveys*, *46*(2), 1–40.

Saghafian, M., Laumann, K., Akhtar, R. S., & Skogstad, M. R. (2020). The evaluation of virtual reality fire extinguisher training. *Frontiers in Psychology, 11*(10), Article 593466. https://doi.org/10.3389/fpsyg.2020.593466

Turner, A., Berry, J., & Holliman, N. (2011, February). *Can the perception of depth in stereoscopic images be influenced by 3D sound?* Paper presented at Stereoscopic Displays and Applications XXII. San Francisco Airport, California, USA: SPIE. https://doi.org/10.1117/12.871960

Wu, B., Ooi, T. L., & He, Z. J. (2004). Perceiving distance accurately by a directional process of integrating ground information. *Nature, 428*(6978), 73–77.

Zahorik, P., Brungart, D. S., & Bronkhorst, A. W. (2005). Auditory distance perception in humans: A summary of past and present research. *Acta Acustica United with Acustica*, *91*(3), 409–420.

Effect of auditory stimulus on distance compression in virtual reality

HU Xiaofei, WANG Jiawei, LIU Hanyu, SONG Xiaolei

(*School of Psychology, Shaanxi Normal University;*
Shaanxi Key Laboratory of Behavior and Cognitive Neuroscience, Xi'an 710000, *China*)

Abstract

Distance compression in virtual reality (VR), which will lead to a distortion of fine manipulation in practical application, depicts that people tend to underestimate the spatial distance of visual stimuli in virtual environments. The apparent perceived differences between virtual and real environments break the immersive experiences and lower users' acceptance. Therefore, it is crucial to ameliorate the distance compression to increase the fidelity and ultimately promote the wider application of VR. Capitalizing on the fact that distance compression is a multiple modality phenomenon and occurs for auditory and visual stimuli, researchers reported that the distance judgment in VR would get more accurate when the positions of auditory and visual stimuli were incongruent. However, it is unclear to what extent the incongruency is to get effective amelioration. In this study, we aimed to completely examine the effect of the auditory stimulus on distance compression in VR. We presumed that the larger the incongruency was, the better amelioration obtained.

We used the HTC Vive Pro to render the virtual environment and the build-in headphone to present auditory stimulus. Thirty participants were recruited to perform a distance judgment task. We first controlled the presence or absence of the auditory stimulus. We also varied the egocentric distance of visual stimulus (3 m, 4 m, 5 m). Then, we controlled the incongruency of the audio-visual condition, that is, the exocentric distance between auditory and visual stimuli (0.5 m, 1 m, 1.5 m, 2 m). Each block consisted of 30 consecutive trials, wherein the reference visual stimulus was presented at the beginning 5 s. Participants were asked to judge whether the following adjust stimulus was nearer or farther than the reference stimulus. The egocentric distance would be adjusted according to the response of the prior trial. The egocentric distance of the adjust stimulus in the last trial was referred to as the ultimate distance judgment for the initial reference stimulus.

We found that the auditory stimulus affected the distance compression in two ways. First, the distance compression under the audio-visual condition was smaller than that under the visual-only condition ($F(1, 29) = 4.05$, $p = 0.054$, $\eta_p^2 = 0.12$), especially when the egocentric distance of the visual stimulus was large (4 m: difference = 1.8%, $t = -1.59$, $df = 29$, $p = 0.062$; 5 m: difference = 1.6%, $t = -1.85$, $df = 29$, $p = 0.037$). Second, we formulated the relationship between the exocentric distance between auditory and visual stimuli and the distance compression rate, which was calculated by subtracting the perceived egocentric distance from the physical egocentric distance and then dividing it by the physical egocentric distance.: distance compression rate = −0.024 × exocentric distance + 0.056. The slope was significant ($p = 0.008$), indicating the distance compression rate was negatively correlated with the exocentric distance and could be ameliorated at a pace of 2.4% for every 1 m. The adjusted R^2 was 90.7%.

We reported the effect of auditory stimulus on the distance compression in VR. Based on our results, we highly recommended presenting the auditory and visual stimuli simultaneously in the time domain and a minimum of 1 m apart in the space domain to ameliorate the distance compression in VR.

Keywords virtual reality, distance compression, auditory and visual stimuli, egocentric distance, exocentric distance